US008135068B1

(12) **United States Patent**
Alvarez et al.

(10) **Patent No.: US 8,135,068 B1**
(45) **Date of Patent: Mar. 13, 2012**

(54) **METHOD AND/OR ARCHITECTURE FOR MOTION ESTIMATION USING INTEGRATED INFORMATION FROM CAMERA ISP**

(75) Inventors: **José R. Alvarez**, Sunnyvale, CA (US); **Guy Cote**, San Jose, CA (US)

(73) Assignee: **Maxim Integrated Products, Inc.**, Sunnyvale, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1294 days.

(21) Appl. No.: **11/769,023**

(22) Filed: **Jun. 27, 2007**

Related U.S. Application Data

(63) Continuation-in-part of application No. 11/410,729, filed on Apr. 25, 2006, now Pat. No. 7,859,574.

(60) Provisional application No. 60/700,789, filed on Jul. 19, 2005.

(51) **Int. Cl.**
H04N 7/12 (2006.01)
G06K 9/36 (2006.01)

(52) **U.S. Cl.** **375/240.16**; 375/240.12; 375/240.13; 375/240.14; 375/240.15; 382/232; 382/233; 382/234

(58) **Field of Classification Search** 375/240.16
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,517,583	A	5/1996	Horiuchi	
5,550,595	A	8/1996	Hannah	
5,619,265	A	4/1997	Suzuki	
5,699,113	A	12/1997	Ohiwa	
6,018,362	A	1/2000	Suzuki	
6,031,939	A	2/2000	Gilbert	
6,046,769	A	4/2000	Ikeda	
6,407,773	B1	6/2002	Ikeda	
6,498,621	B1	12/2002	Ikeda	
7,489,344	B2 *	2/2009	Fukuoka et al.	348/222.1
2006/0204115	A1 *	9/2006	Burazerovic	382/239

OTHER PUBLICATIONS

US Office Action for related copending U.S. Appl. No. 11/410,729; dated Jun. 3, 2010.

US Office Action for related copending U.S. Appl. No. 11/410,729; dated Nov. 18, 2009.

US Office Action for related issued patent No. 7,733,380; dated Aug. 5, 2009.

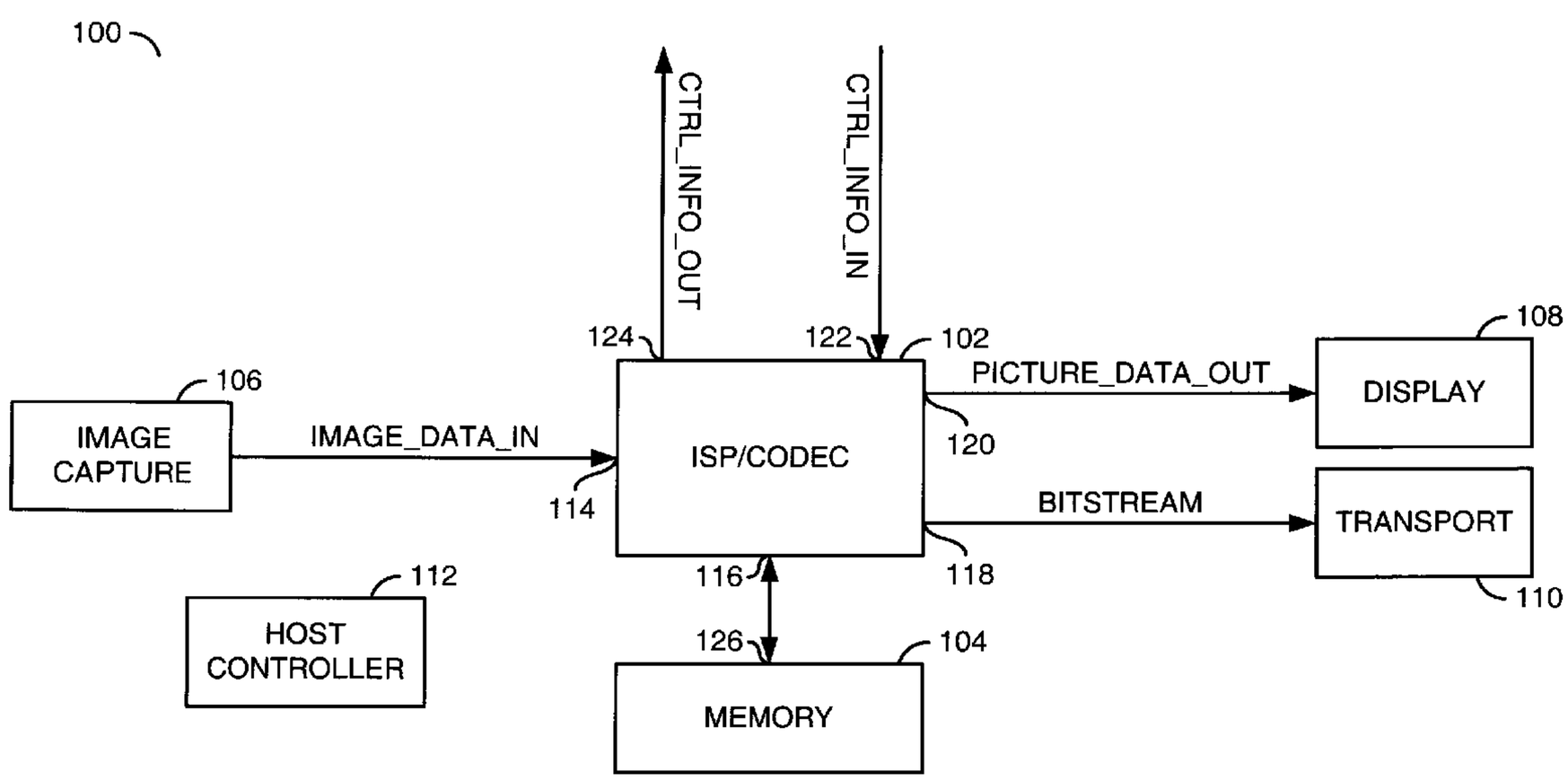
* cited by examiner

Primary Examiner — Longbit Chai
(74) *Attorney, Agent, or Firm* — North Weber & Baugh LLP

(57) **ABSTRACT**

A camera comprising a first circuit and a second circuit. The first circuit may be configured to perform image signal processing using encoding related information. The second circuit may be configured to encode image data using image signal processing related information. The first circuit may be further configured to pass the image signal processing related information to the second circuit. The second circuit may be further configured to pass the encoding related information to the first circuit. The second circuit may be further configured to modify one or more motion estimation processes based upon the information from the first circuit.

23 Claims, 8 Drawing Sheets



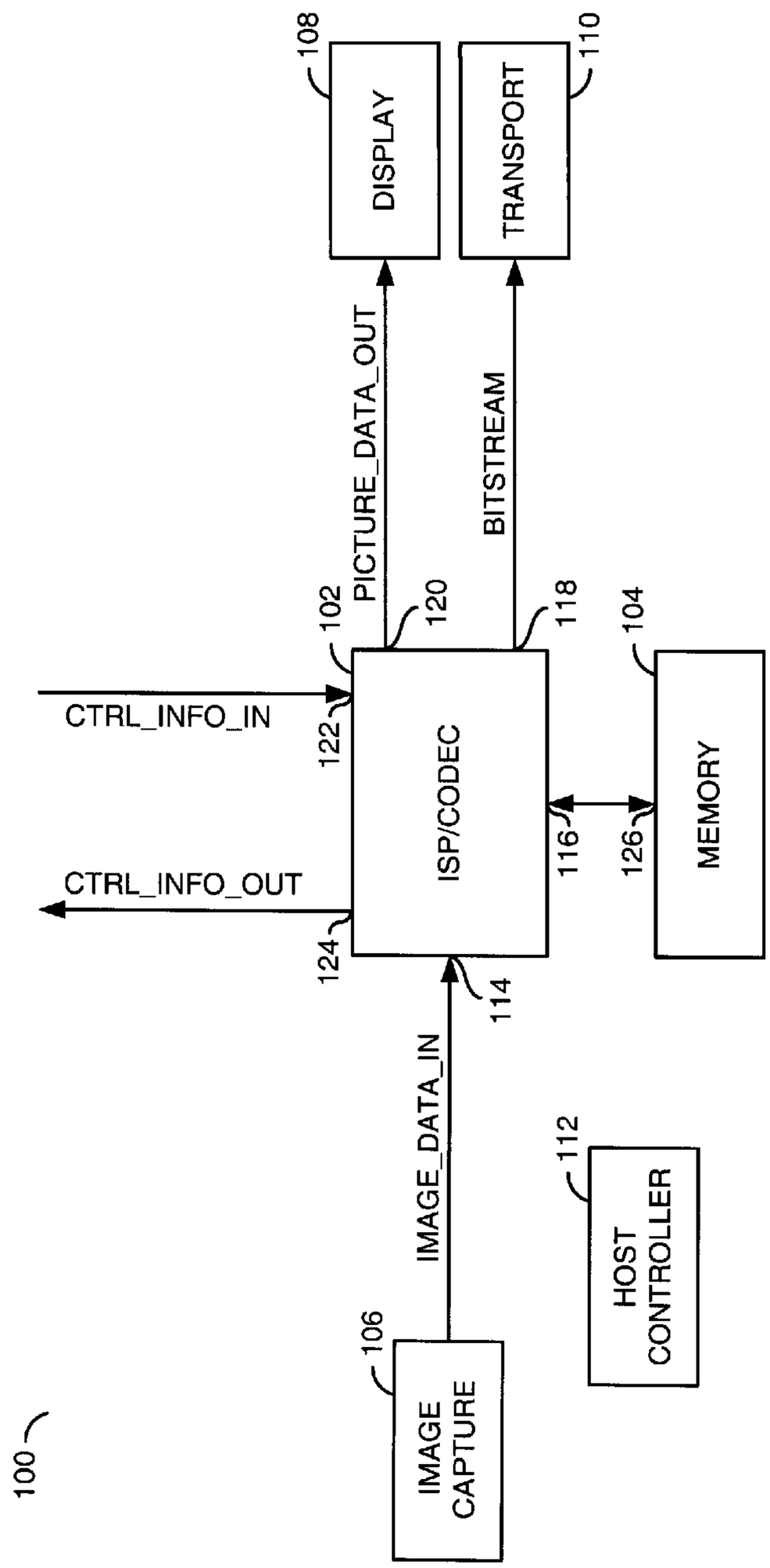

100
IMAGE CAPTURE
106
IMAGE_DATA_IN
114
HOST CONTROLLER
112
CTRL_INFO_OUT
124
CTRL_INFO_IN
122
ISP/CODEC
102
PICTURE_DATA_OUT
120
BITSTREAM
118
116
126
MEMORY
104
DISPLAY
108
TRANSPORT
110

FIG. 1

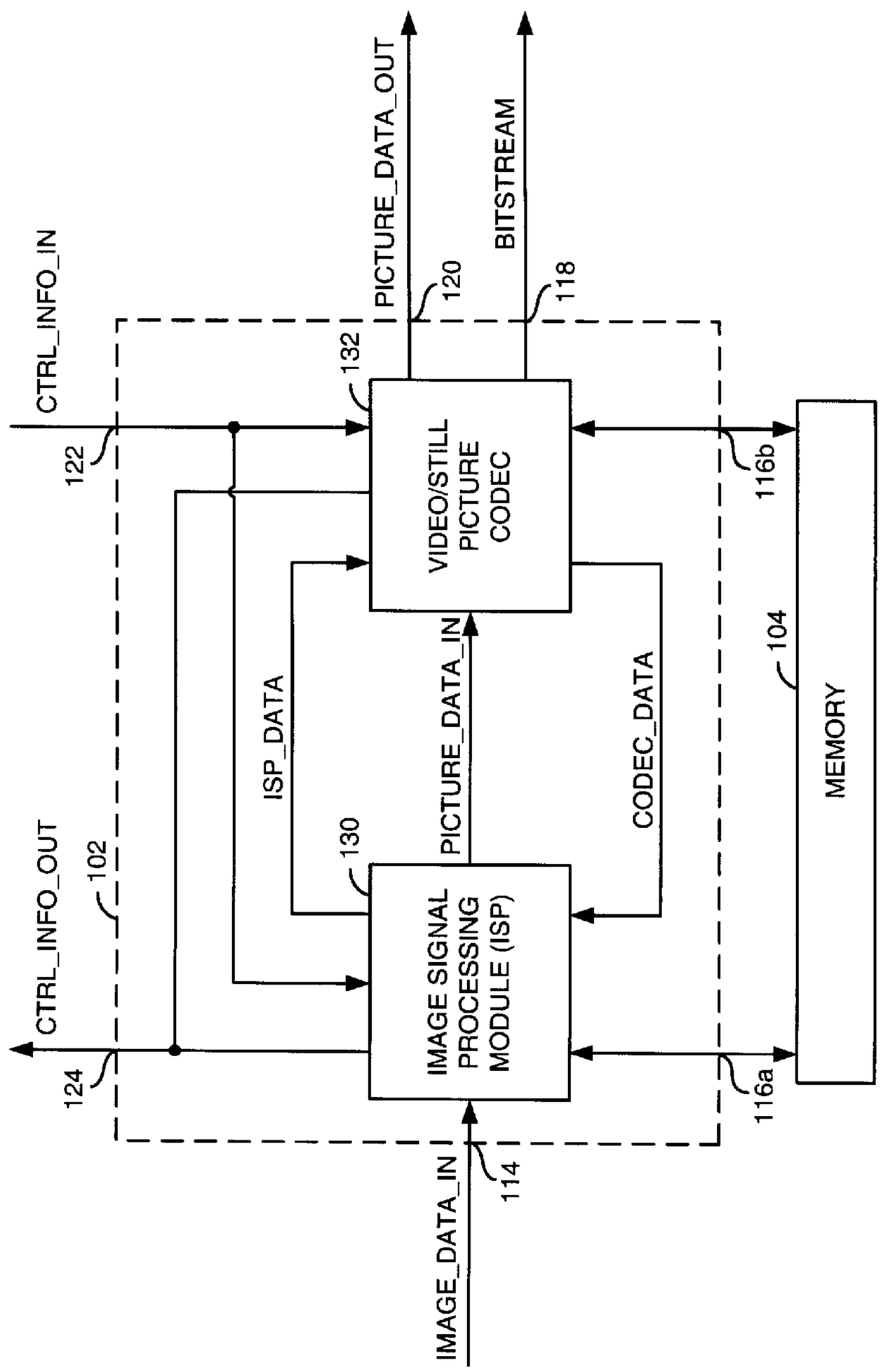

CTRL_INFO_IN
122
PICTURE_DATA_OUT
120
BITSTREAM
118
132
VIDEO/STILL PICTURE CODEC
116b
104
MEMORY
ISP_DATA
PICTURE_DATA_IN
CODEC_DATA
130
IMAGE SIGNAL PROCESSING MODULE (ISP)
CTRL_INFO_OUT
102
124
116a
114
IMAGE_DATA_IN

FIG. 2

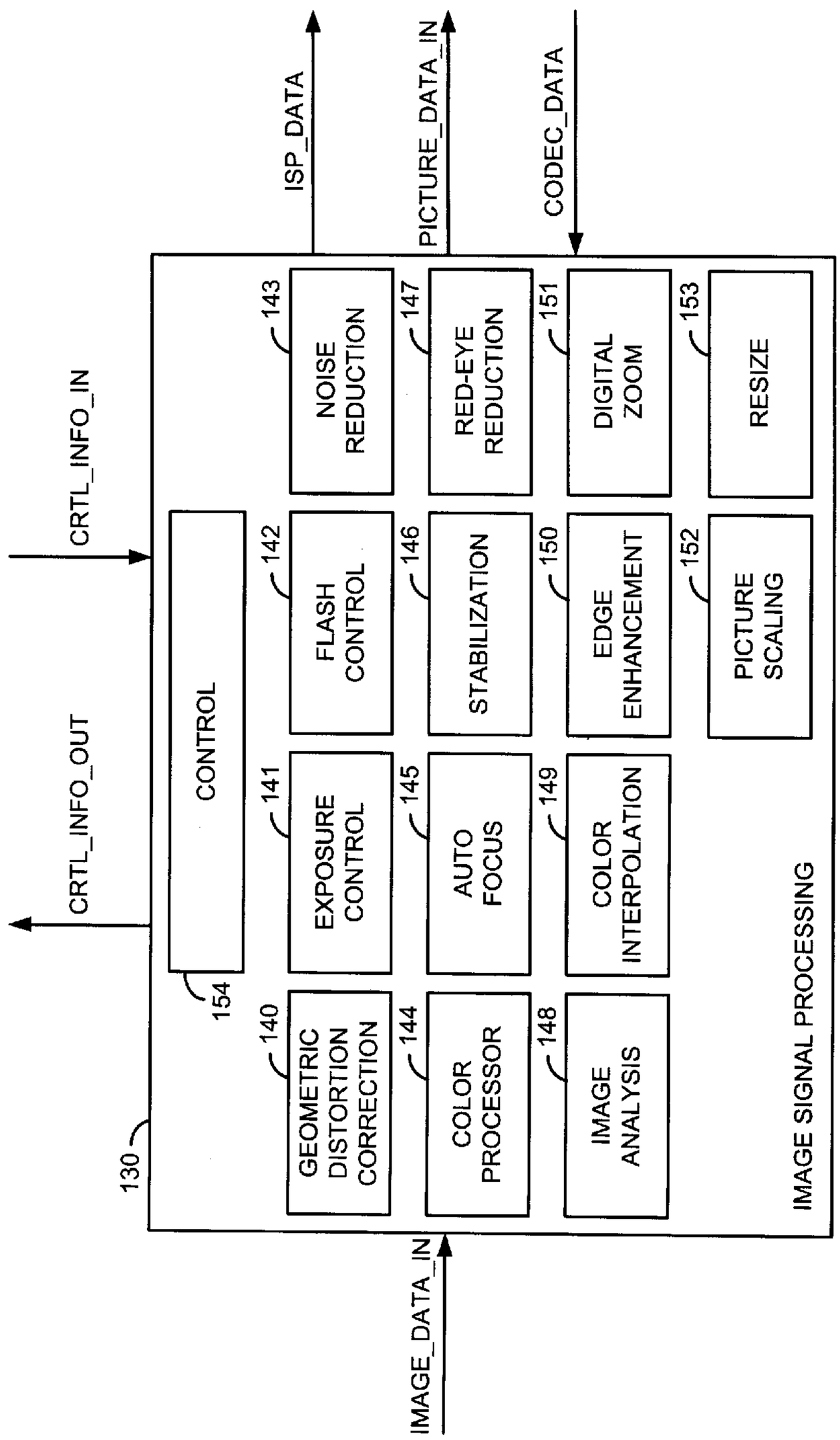

ISP_DATA
PICTURE_DATA_IN
CODEC_DATA
CRTL_INFO_IN
CRTL_INFO_OUT
130
CONTROL
154
NOISE REDUCTION
143
RED-EYE REDUCTION
147
DIGITAL ZOOM
151
RESIZE
153
FLASH CONTROL
142
STABILIZATION
146
EDGE ENHANCEMENT
150
PICTURE SCALING
152
EXPOSURE CONTROL
141
AUTO FOCUS
145
COLOR INTERPOLATION
149
GEOMETRIC DISTORTION CORRECTION
140
COLOR PROCESSOR
144
IMAGE ANALYSIS
148
IMAGE SIGNAL PROCESSING
IMAGE_DATA_IN

FIG. 3

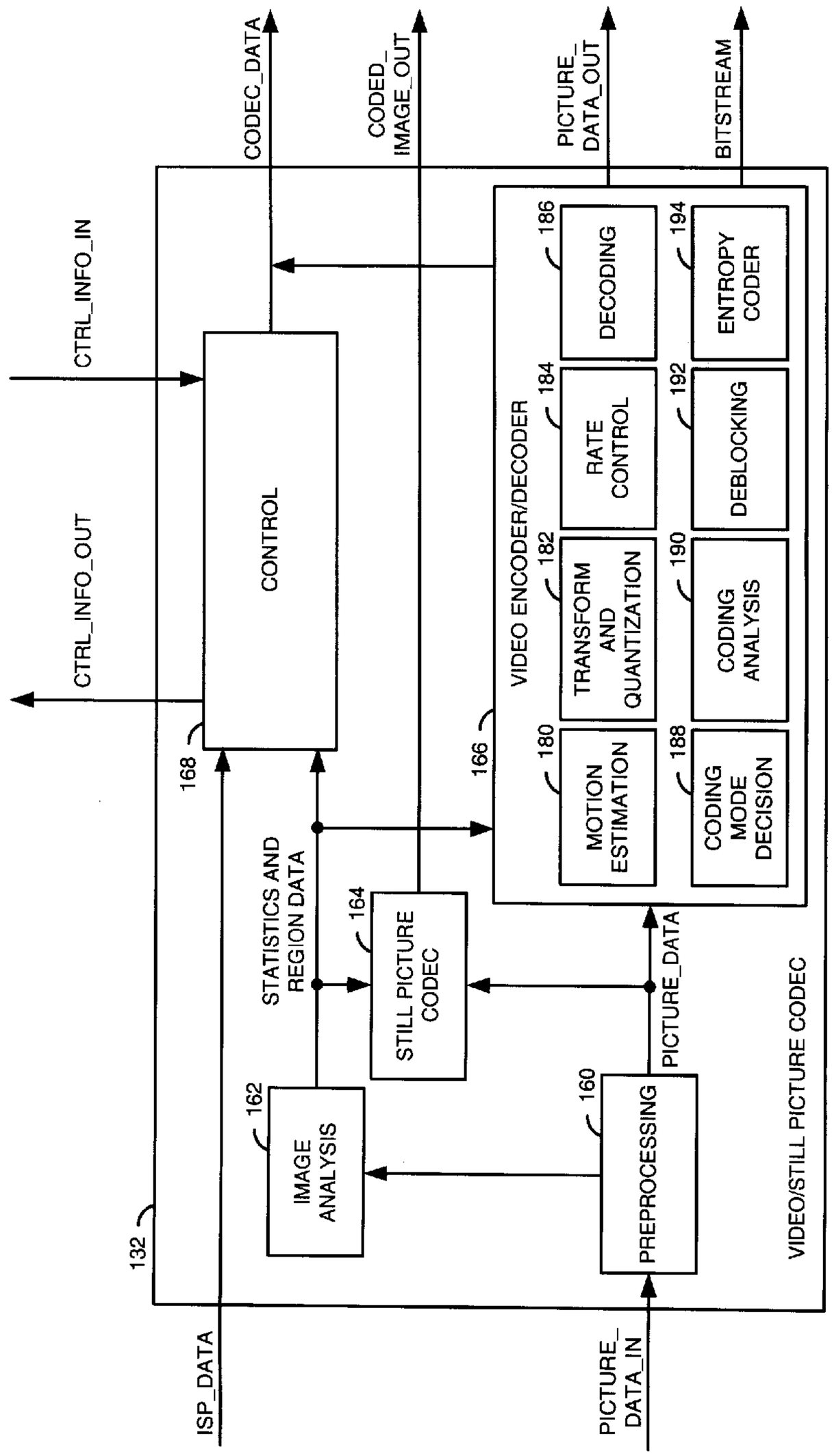

CODEC_DATA
CODED_ IMAGE_OUT
PICTURE_ DATA_OUT
BITSTREAM
CTRL_INFO_IN
CTRL_INFO_OUT
CONTROL
168
VIDEO ENCODER/DECODER
166
DECODING
186
ENTROPY CODER
194
RATE CONTROL
184
DEBLOCKING
192
TRANSFORM AND QUANTIZATION
182
CODING ANALYSIS
190
MOTION ESTIMATION
180
CODING MODE DECISION
188
STATISTICS AND REGION DATA
STILL PICTURE CODEC
164
PICTURE_DATA
IMAGE ANALYSIS
162
PREPROCESSING
160
VIDEO/STILL PICTURE CODEC
132
ISP_DATA
PICTURE_ DATA_IN

FIG. 4

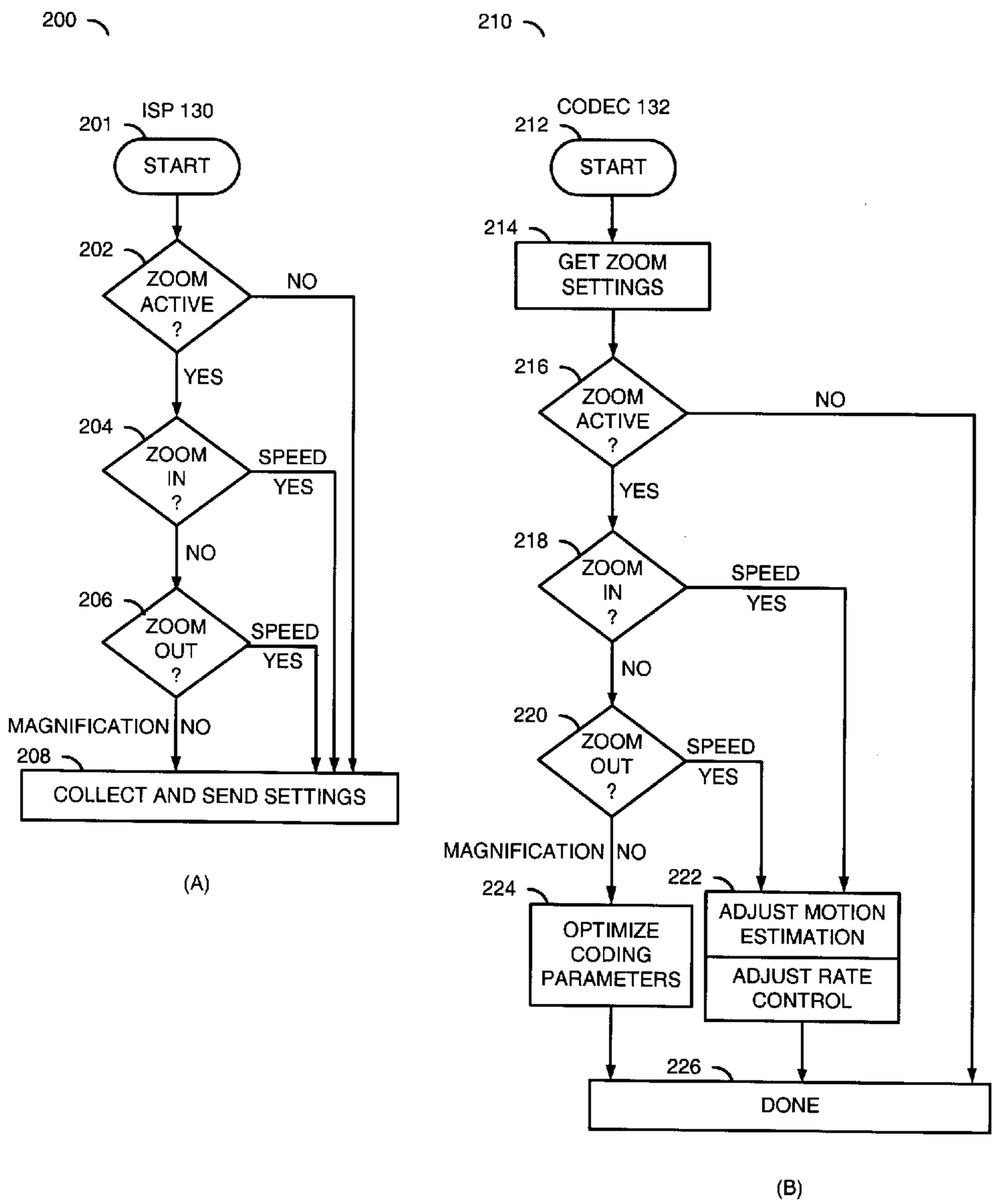

200
ISP 130
201
START
202
ZOOM ACTIVE ?
NO
YES
204
ZOOM IN ?
SPEED YES
NO
206
ZOOM OUT ?
SPEED YES
MAGNIFICATION NO
208
COLLECT AND SEND SETTINGS
(A)
210
CODEC 132
212
START
214
GET ZOOM SETTINGS
216
ZOOM ACTIVE ?
NO
YES
218
ZOOM IN ?
SPEED YES
NO
220
ZOOM OUT ?
SPEED YES
MAGNIFICATION NO
224
OPTIMIZE CODING PARAMETERS
222
ADJUST MOTION ESTIMATION
ADJUST RATE CONTROL
226
DONE
(B)

FIG. 5

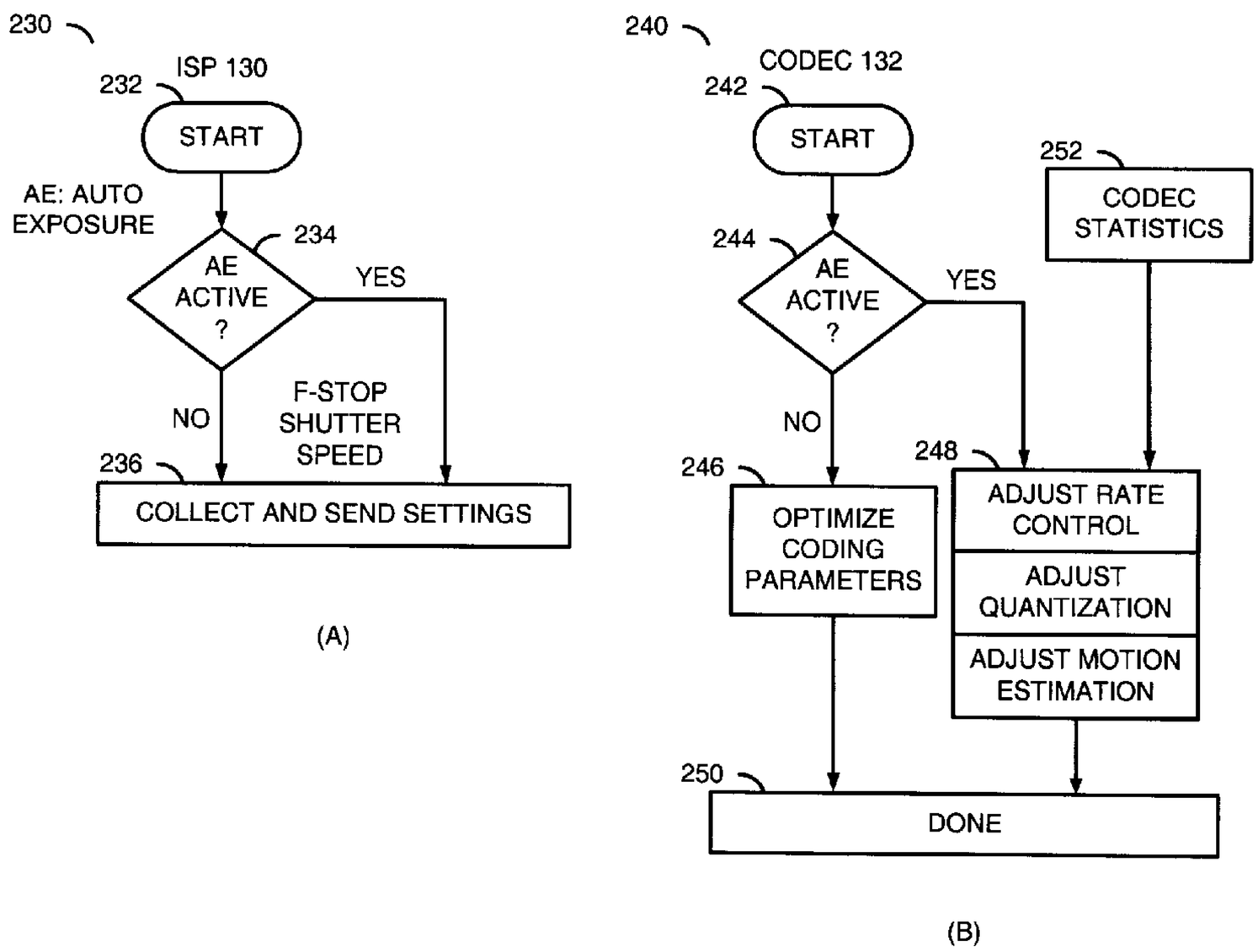

230
ISP 130
232
START
AE: AUTO EXPOSURE
234
AE ACTIVE ?
YES
NO
F-STOP SHUTTER SPEED
236
COLLECT AND SEND SETTINGS
(A)
240
CODEC 132
242
START
252
CODEC STATISTICS
244
AE ACTIVE ?
YES
NO
246
OPTIMIZE CODING PARAMETERS
248
ADJUST RATE CONTROL
ADJUST QUANTIZATION
ADJUST MOTION ESTIMATION
250
DONE
(B)

FIG. 6

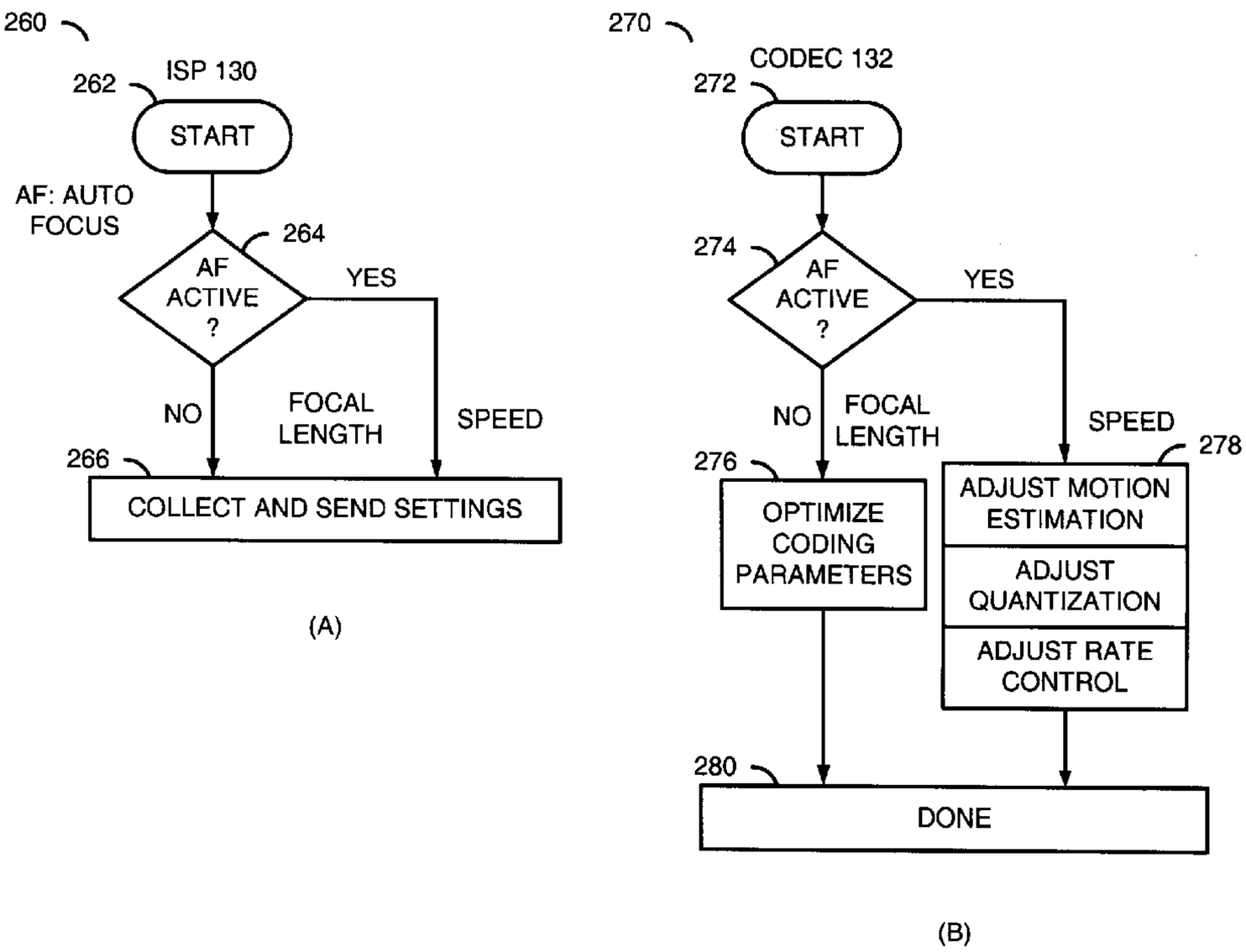

260
ISP 130
262
START
AF: AUTO FOCUS
264
AF ACTIVE ?
YES
NO
FOCAL LENGTH
SPEED
266
COLLECT AND SEND SETTINGS
(A)
270
CODEC 132
272
START
274
AF ACTIVE ?
YES
NO
FOCAL LENGTH
SPEED
276
278
OPTIMIZE CODING PARAMETERS
ADJUST MOTION ESTIMATION
ADJUST QUANTIZATION
ADJUST RATE CONTROL
280
DONE
(B)

FIG. 7

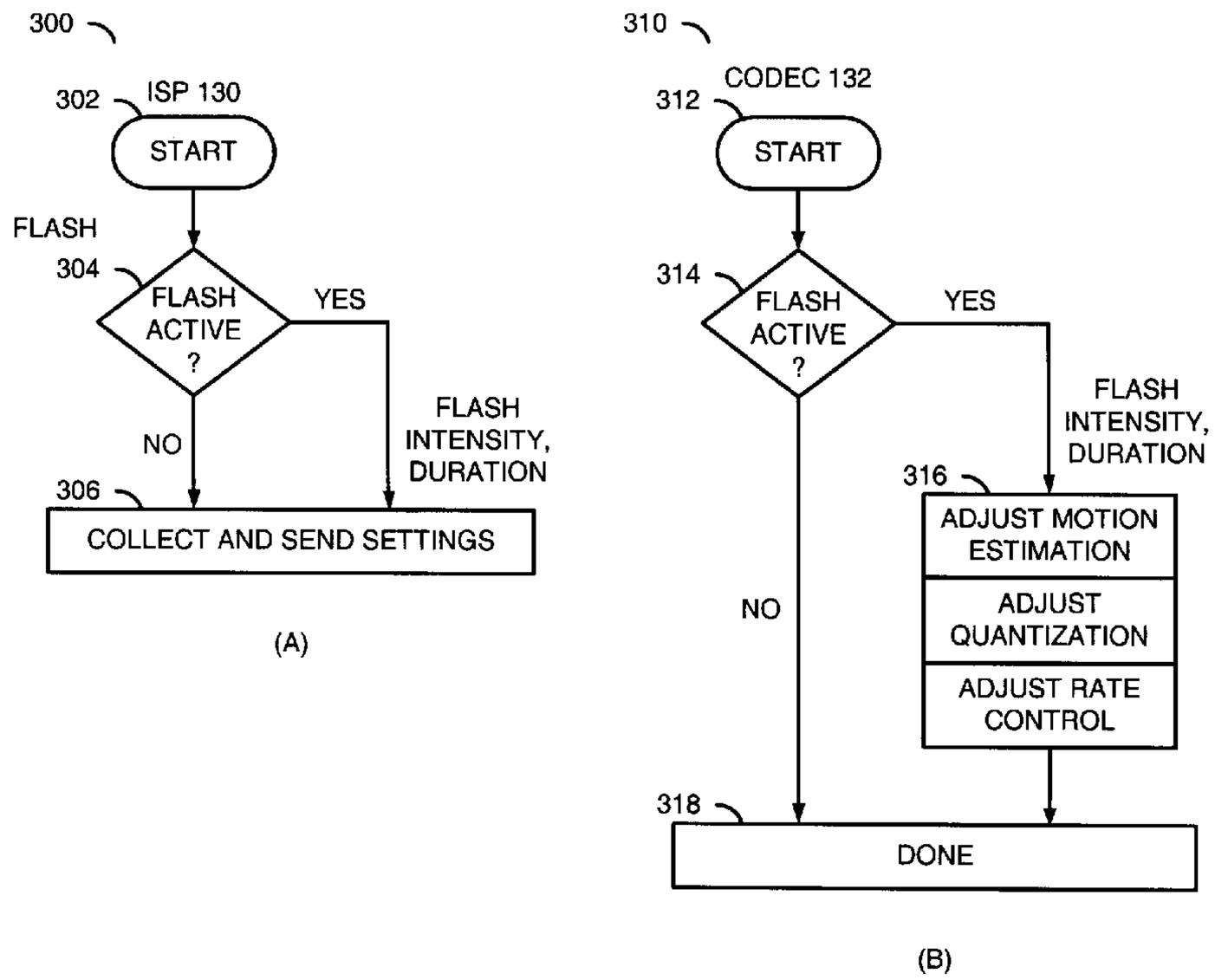

300
302
ISP 130
START
FLASH
304
FLASH ACTIVE ?
YES
NO
FLASH INTENSITY, DURATION
306
COLLECT AND SEND SETTINGS
(A)
310
312
CODEC 132
START
314
FLASH ACTIVE ?
YES
NO
FLASH INTENSITY, DURATION
316
ADJUST MOTION ESTIMATION
ADJUST QUANTIZATION
ADJUST RATE CONTROL
318
DONE
(B)

FIG. 8

METHOD AND/OR ARCHITECTURE FOR MOTION ESTIMATION USING INTEGRATED INFORMATION FROM CAMERA ISP

The present application is a Continuation-In-Part of U.S. application Ser. No. 11/410,729, filed Apr. 25, 2006 now U.S. Pat. No. 7,859,574, which claims the benefit of U.S. Provisional Application No. 60/700,789, filed Jul. 19, 2005, which are each hereby incorporated by reference in their entirety.

The present application may relate to U.S. application Ser. No. 11/564,483, filed Nov. 29, 2006, and U.S. application Ser. No. 11/675,715, filed Feb. 16, 2007, which are hereby incorporated by reference in their entirety.

FIELD OF THE INVENTION

The present invention relates to video processing generally and, more particularly, to a method and/or architecture for motion estimation using integrated information from camera image signal processor (ISP).

BACKGROUND OF THE INVENTION

Economies of scale in non-volatile storage and electronics in general are making it possible to create commercially attractive devices capable of high quality digital still picture photography and high quality digital video recording in the same mobile device. However, typical devices are either efficient at processing digital still pictures (Digital Still Cameras, DSC) or digital video streams (Digital Video Cameras, DVC), but not both.

SUMMARY OF THE INVENTION

The present invention concerns a camera comprising a first circuit and a second circuit. The first circuit may be configured to perform image signal processing using encoding related information. The second circuit may be configured to encode image data using image signal processing related information. The first circuit may be further configured to pass the image signal processing related information to the second circuit. The second circuit may be further configured to pass the encoding related information to the first circuit. The second circuit may be further configured to modify one or more motion estimation processes based upon the information from the first circuit.

The objects, features and advantages of the present invention include providing a method and/or architecture for motion estimation using integrated information from camera ISP that may (i) take advantage of processing and system synergies, (ii) provide an efficient digital still/video camera, (iii) provide a low cost digital still/video camera, (iv) provide improved performance over separate camera implementations and/or (v) provide improved image signal processing and encoding/decoding by exchanging information between image signal processor (ISP) and Codec functions.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other objects, features and advantages of the present invention will be apparent from the following detailed description and the appended claims and drawings in which:

FIG. **1** is a block diagram illustrating an ISP/Codec in accordance with the present invention;

FIG. **2** is a more detailed block diagram illustrating connections between an image processing module and an encoding/decoding module of the ISP/Codec of FIG. **1**;

FIG. **3** is a block diagram illustrating functions of the image signal processing module of FIG. **2**;

FIG. **4** is a block diagram illustrating functions of the coder/decoder module of FIG. **2**;

FIGS. **5**(A-B) are flow diagrams illustrating an example of a zoom operation in accordance with a preferred embodiment of the present invention;

FIGS. **6**(A-B) are flow diagrams illustrating an example of an auto exposure operation in accordance with a preferred embodiment of the present invention;

FIGS. **7**(A-B) are flow diagrams illustrating an example of an auto focus operation in accordance with a preferred embodiment of the present invention; and

FIGS. **8**(A-B) are flow diagrams illustrating an example of a flash operation in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

A commercial mobile/portable system that is as efficient as a digital still camera (DSC) and a digital video camera (DVC) involves the use of specific hardware/software. However, if the hardware/software systems for the DSC and the DVC are developed independently, allowing minimal interaction with each other, processing and system synergies between the two systems are not fully utilized. The present invention generally provides a system for taking advantage of such synergies to produce an efficient and low cost combination DSC/DVC with better performance than separate implementations.

Referring to FIG. **1**, a block diagram is shown illustrating a still and video picture processing system **100** in accordance with the present invention. In one example, the system **100** may include a block (or circuit) **102**, a block (or circuit) **104**, a block (or circuit) **106**, a block (or circuit) **108**, a block (or circuit) **110** and a block (or circuit) **112**. The block **102** may comprise an image signal (or sensor) processing (ISP) function (or circuit) and an encoding (or encoding/decoding (Codec)) function (or circuit). In one example, the block **102** may be implemented as a single integrated circuit chip.

The block **104** may be implemented, in one example, as a memory. However, other appropriate storage media or devices may be implemented accordingly to meet the design criteria of a particular implementation. The block **106** may be implemented, in one example, as an image capture module. The block **108** may be implemented, in one example, as an image display module. The block **110** may be implemented, in one example, as a data transport module. The block **112** may be implemented, in one example, as a host controller.

In one example, the block **102** may have an input **114** that may receive a signal (e.g., IMAGE_DATA_IN), an input/output **116** that may be configured to connect the block **102** with the block **104**, an output **118** that may present a signal (e.g., BITSTREAM), an output **120** that may present a signal (e.g., PICTURE_DATA_OUT), an input **122** that may receive a signal or signals (e.g., CTRL_INFO_IN) and an output **124** that may present a signal or signals (e.g., CTRL_INFO_OUT). The signal BITSTREAM may be implemented, in one example, as a compressed bitstream. In one example, the signal BITSTREAM may comprise data compressed in compliance with a standard or proprietary scheme (e.g., JPEG, MPEG, H.264, VC1, VC9, etc.). The signal IMAGE_DATA_IN may be implemented, in one example, as a still picture and/or video input signal. The signal PICTURE_DATA_OUT may be implemented, in one example, as a still picture and/or video output signal. The signal(s) CTRL_INFO_IN may comprise a number of types

of information (e.g., capture control information, stabilization parameters, flash settings, images, active parameters, exposure settings, noise reduction settings, motion vectors and/or capture mode information). The signal(s) CTRL_INFO_OUT may comprise a number of types of data (e.g., capture parameters, images, rate control information, statistical information, exposure settings, correlation analysis, etc.).

The block **104** may have an input/output **126** that may be configured to couple (or interface) the block **104** to the block **102**. The block **106** may be configured to capture image data, either still or moving for presentation to the block **102**. For example, the block **106** may be implemented, in one example, as a lens and sensor (e.g., CCD, etc.). The block **108** may be implemented in one example, as any of a number of types of image display devices (e.g., CRT, LCD, plasma, etc.). The block **110** may be implemented as any of a number of transport and/or storage systems (e.g., CD, DVD, PVR, Internet, satellite, broadcast, etc.). The block **112** may be configured to control various operations of the system **100**. Although interconnections to the other blocks are not shown, the block **112** may be configured to provide control signals and receive control data from each or all of the various blocks of the system **100**, as would be recognized by one of ordinary skill in the art.

External images may be prepared for input to the block **102** by the block **106**. For example, the external images may be converted from an analog domain to a digital domain. In one example, the images may be converted by means of an image sensor and associated analog and digital processing circuits. In one example, the images may be converted by the block **106** from the analog domain into a format that is conducive to further digital processing. For example, the block **106** may be configured to convert and process data into a two-dimensional sampling matrix. In one example, the process used may include a generally known technique of demosaicing (or color interpolation). Conventional techniques may be used for demosaicing.

The blocks **108** and **110** are generally illustrative of output processes. The block **108** may be configured to display decoded bit stream information (e.g., on a monitor). The block **110** may be configured to transport the encoded bit stream (e.g., as an MPEG-2 transport stream, etc.). The block **112** is generally configured to oversee the entire process and coordinate control and data processing among all the various modules in the system **100**.

Referring to FIG. **2**, a more detailed block diagram is shown illustrating an example implementation of the block **102** of FIG. **1**. In one example, the block **102** may comprise a block (or circuit) **130** and a block (or circuit) **132**. The block **130** may be implemented, in one example, as an image signal (or sensor) processing (ISP) module. In one example, the block **130** may be implemented using a digital signal processor (DSP). The block **132** may be implemented, in one example, as a video/still picture encoder/decoder (Codec). In other examples, the block **132** may be implemented as a video/still picture encoder, a video encoder or a still picture encoder. In one example, the block **104** may be implemented as a multi-port memory. Although the block **104** is generally illustrated as a separate block from the block **102**, one of skill in the relevant art would recognize that the function of the block **104** may be integrated with the block **102** and/or each of the blocks **130** and **132**.

The block **130** may have a first input that may receive the signal IMAGE_DATA_IN, a first output that may present information (or data) as the signal CTRL_INFO_OUT, a second input that may receive the signal CTRL_INFO_IN, a second output that may present a signal (e.g., ISP_DATA), a third output that may present a signal (e.g., PICTURE_DATA_IN), a third input that may receive a signal (e.g., CODEC_DATA), and an input/output that may connect, in one example, to a first port of the block **104**. The block **132** may have a first input that may receive the signal PICTURE_DATA_IN, a second input that may receive the signal ISP_DATA, a first output that may present information (or data) as the signal CTRL_INFO_OUT, a third input that may receive the signal CTRL_INFO_IN, a second output that may present the signal PICTURE_DATA_OUT, a third output that may present the signal BITSTREAM, a fourth output that may present the signal CODEC_DATA and an input/output that may be coupled, in one example, to a second port of the block **104**.

The block **130** may be configured to convert raw image data to picture data amenable to compression. The block **130** may be configured to perform several processing steps for creating a picture that is both true to the image in the real world, and properly realized by the processing circuits, taking into account limitations of the technology used for the process. In one example, the limitations may include (i) finite dynamic ranges and temporal response in image sensor technology, (ii) inherent distortions in lenses, (iii) inherent noise in the capture process, (iv) errors and inconsistencies in transducers used to control exposure, focus and color representation and/or (v) processing limitations in signal processing techniques.

Properly exposed and corrected picture data may be presented to the block **132** for further processing and compression. The block **132** may use, for example, hybrid entropy coding techniques such as MPEG, H.264, and/or JPEG to reduce the amount of picture data for storage and/or transmission. In one example, decoded (or decompressed) data may be presented for display on an external device. As used herein, the terms MPEG, H.264 and JPEG are used to generally refer to a variety of coding standards and/or proprietary schemes that may include, but are not limited to, MPEG-1, MPEG-2, MPEG-4, H.264, JPEG, JPEG-2000, etc.

Data (or image signal processing related information) generated from processes in the block **130** may be transmitted via the signal ISP_DATA to the block **132**. Similarly, data (or coding/decoding related information) generated from processes in the block **132** may be transmitted via the signal CODEC_DATA to the block **130**. In one example, the data may be transmitted using a communication protocol that allows for control and information to be interchanged between the modules. The information transmitted from the block **130** to the block **132** and from the block **132** to the block **130** is generally implemented for improving the quality of processing in both blocks (or systems). For example, the processes of the block **130** and the block **132** may be configured to modify (or adapt, change, optimize, etc.) operations based upon the data received from the block **132** and the block **130**, respectively (described in more detail below in connection with FIGS. **5-8**). The present invention may produce still pictures and video sequences of better quality than if the still pictures or the sequences had been processed independently by separate modules.

In general, processing improvements in the block **130** may be derived from data sent from the block **132**. The improvements are generally independent of individual improvements in each of the processing steps involved in a normal image signal processor. Advancements in each processing operation in the block **130** may advance the performance of the system as a whole, but benefit may still be realized from data provided by the block **132** in accordance with the present inven-

tion. The data sent from the block **132** to the block **130** may comprise control data, statistical data and/or global and/or region-based spatio-temporal data derived from analysis of the image and video sequence data.

Similarly, video coding improvements by the block **132** may be realized by using data received from the block **130** in accordance with the present invention. The improvements are generally independent of individual advancements in each module in the block **132**. The following sections generally describe examples of individual improvements that may be achieved by generating, transferring and using the data transferred between the block **130** and the block **132** in accordance with the present invention. The examples are presented as an illustration of, and not a limitation on, the improvements that may be realized using the present invention.

Referring to FIG. **3**, a more detailed block diagram illustrating an example of the image signal (or image sensor) processing module **130** of FIG. **2** is shown. In one example, the block **130** may comprise a block (or circuit) **140**, a block (or circuit) **141**, a block (or circuit) **142**, a block (or circuit) **143**, a block (or circuit) **144**, a block (or circuit) **145**, a block (or circuit) **146**, a block (or circuit) **147**, a block (or circuit) **148**, a block (or circuit **149**, a block (or circuit) **150**, a block (or circuit) **151**, a block (or circuit) **152**, a block (or circuit) **153** and a block (or circuit) **154**. The block **140** may be implemented, in one example, as a geometric distortion correction block. The block **141** may be implemented, in one example, as an exposure control module. The block **142** may be implemented, in one example, as a flash control module. The block **143** may be implemented, in one example, as a noise reduction module. The block **144** may be implemented, in one example, as a color processor module. The block **145** may be implemented, in one example, as a auto focus module. The block **146** may be implemented, in one example, as a stabilization module. The block **147** may be implemented, in one example, as a red-eye reduction module. The block **148** may be implemented, in one example, as an image analysis module. The block **149** may be implemented, in one example, as a color interpolation block. The block **150** may be implemented, in one example, as an edge enhancement block. The block **151** may be implemented, in one example, as a digital zoom block. The block **152** may be implemented, in one example, as a picture scaling block. The block **153** may be implemented, in one example, as a resizing block. The block **154** may be implemented, in one example, as a control block.

The blocks **140-154** may be configured to perform a series of operations on the captured image data (e.g., received via the signal IMAGE_DATA_IN) in order to improve quality and compression of the data. In one example, the blocks **140-154** may be implemented by modifying conventional techniques in accordance with the teaching of the present disclosure. The blocks **140-154** may be implemented, in one example, using a digital signal processor (DSP). For example, a specialized processor may be implemented containing dedicated hardware/software processing modules to achieve each operation illustrated in FIG. **3**.

In general, real world images may be captured and converted from the analog domain to the digital domain. The block **140** may provide for correcting geometric distortions in the optical system of the image capture block **106** (e.g., distortion and shading compensation). The block **141** may provide automatic exposure functions to ensure that image illuminance/color ranges are properly captured by the imaging sensor of the block **106**. For low light situations, the block **143** may provide noise reduction techniques that may be implemented to improve the signal-to-noise ratio (SNR) while the block **142** may provide automatic settings of the external flash that may be used to improve scene illumination. The block **144** may perform color processing (e.g., white balance, color correction, gamma correction, color conversion, etc.). The block **145** may be configured to automatically adjust a focal plane of the lens to ensure that the subject is in focus.

Under low lighting conditions with increased exposure time, the possibility of image blur may be reduced by the block **146** implementing an electronic image stabilization process configured to operate in concert with other processing modules in the block **130**. Furthermore, in low light environments, the use of an external flash may cause a "red-eye" condition in some subjects. The block **147** may implement automatic red-eye reduction techniques that may be included in the system **100**.

The image and picture quality provided by the block **130** may be further improved through image analysis performed in the block **148**, color interpolation processes performed by the block **149**, edge enhancement processes implemented in the block **150** and/or digital zoom operations performed by the block **151**. Image and picture size and/or aspect ratio may be adjusted by the block **152** and the block **153**. In general, the operation of the blocks **140-153** may be controlled via the control block **154**. The block **130** is generally capable of producing quality images and pictures that may be further processed by a still picture compression module and/or a video compression module. In one example, the still picture compression module and/or video compression module may be implemented external to the block **130**. For example, a still picture encoder (or encoder/decoder) process and/or an advanced video encoder (AVC), or encoder/decoder (Codec) process may reside in the block **132**.

The signal ISP_DATA may be generated dynamically as the block **130** is optimizing image capture for picture quality improvements. The signal CODEC_DATA sent from the block **132** to the block **130** may comprise control and statistics data and employ a mutually agreed upon transmission protocol that includes specific commands to adjust parameters of the block **130**.

As the pictures are encoded in real time, the block **130** may adjust operations in order to improve performance using statistical and image characteristics data transmitted back from the block **132** (described in more detail below). For example, the block **130** may be configured to modify processes based on data received from the block **132** in order to improve picture quality. In general, all of the image signal processing techniques implemented in the block **130** may benefit from real-time data gathered by the block **132** since image signal processing is generally not optimized to process sequences of pictures but rather still images.

The control information received by the block **130** from the block **132** may be used not only for improving image processing, but also to improve coding efficiency. The ISP settings in the block **130** may be deliberately modified to affect both the still image and moving picture coding processes of the block **132**. Furthermore, many of the techniques described herein may be tuned by a user of the system. For example, an operator may set parameters to achieve a particular quality level in still picture/video capture situations.

The block **141** may be configured to provide automatic exposure processing for the block **130**. By examining the information from light sensors, the block **141** may set exposure time and lens aperture to optimize image capture by the image sensor device. The sensitivity of the sensor is commonly expressed in terms of traditional film sensitivity as dictated by the International Organization for Standardization (ISO). A properly exposed picture generally takes advantage of the dynamic range of the sensor. In some cases, the

processing provided by the block **130** may optimize the dynamic range so that particular areas of the picture are better exposed. For example, particular areas may be deemed more important (e.g., people's faces in shadowed areas).

The block **142** may provide automatic flash control for the block **130**. Under low lighting situations, the block **142** may automatically determine the need for external flash illumination given the chosen (or derived) exposure parameters (e.g., aperture, shutter speed, ISO sensitivity). The block **142** may control the external flash via the signal CTRL_INFO_OUT.

The block **143** may provide noise reduction for the block **130**. Under low lighting conditions or high ISO equivalent capture settings, the possibility of random noise generally increases. The block **143** may reduce random noise to improve picture quality. In typical DSC configurations there is limited use of temporal domain information to reduce noise. Lowering the amount of random noise in pictures increases perceived quality by allowing better compression. Lowering the amount of random noise also allows better capture parameters to be used for each picture (e.g., higher shutter speeds and smaller apertures).

The block **145** may provide automatic focus control/processing for the block **130**. In one example, a normal operation of the automatic focus control may be to hunt for a proper focus point. Such hunting behavior may be modified based on information (e.g., motion vector, etc.) provided by the block **132** in accordance with the present invention. For example, a technique for focusing a camera may implement an auto-focusing mechanism that re-focuses the image until a greatest amount of high frequencies are obtained. For example, the point where the most high frequencies are obtained may be associated with the perfect focus of the point because the image no longer looks blurry. The present invention may use statistics about where a high number of edges are present in an image to control the auto-focusing mechanism. For example, the auto-focusing mechanism may be adjusted to focus on a small portion of the image containing edges to improve the performance of the auto-focus.

In another example, the block **145** may analyze a set of areas of the picture to determine the proper focal length for sharp image capture. In general, the block **130** may be configured to identify the main subject of the picture especially when the aperture is wide (i.e., the depth-of-field is shallow) in order to properly bring into focus the important areas of the picture.

The block **146** may be configured to provide image stabilization. Under low lighting situations or large zoom magnification, the possibility of camera motion (e.g., shake) and, therefore, blurry pictures increases. Large zoom factors generally employ larger apertures and therefore create shallow depth-of-field conditions that call for sharp, continuous focusing. The block **154** generally manages the conflicting focusing demands and modifies the efficacy of the stabilization techniques to improve exposure and/or focus mechanisms.

The block **147** may be configured to perform red-eye reduction for the block **130**. Under low lighting conditions, some subjects will reflect flash illumination from the back of the eye (e.g., the retina) therefore causing the red-eye effect. The block **147** may be configured to determine such situations and adapt/correct the areas where red-eye is most likely to occur. In some implementations, the flash may be pulsed rapidly prior to final picture capture so as to narrow the pupil aperture of a subject.

The block **148** may be configured to derive picture characteristics representing spatial and temporal statistics data (e.g., averages, standard deviations, probability distributions, histograms, etc.). The block **148** may be further configured to collect feature-based image characteristics (e.g., object texture, color, edges, motion, etc.) and integrate the characteristics into regions of interest based on a combination of the picture features.

The majority of image sensors generate more spatial information than allowed by the physical number of photosensitive receptors by sampling each color component with different resolution. For example, a subsampling scheme referred to as Bayer pattern sampling may be used. In the Bayer pattern sampling scheme, the red and blue components are sampled once every four spatial positions, while the green component is sampled twice every four spatial positions. The Bayer pattern sampling scheme uses a two-dimensional sampling pattern illustrated in the following example array:

G	B
R	G

The color interpolation block **149** may be used to generate a complete signal for all color components at all spatial positions through digital interpolation to produce the same number of samples for all the red, green and blue components.

The edge enhancement block **150** may be used to increase the perceived resolution of sampled images. For example, the block **150** may detect the location and strength of object edges in two-dimensions and increase the contrast of the edges with respect to the surrounding picture detail, therefore creating the appearance of more resolution.

The blocks **151**, **152** and **153** are generally related to each other in that the blocks **151**, **152** and **153** may change the resolution of the pictures based upon predefined (or programmable) criteria. The digital zoom block **151** may be used in place of an optical zoom mechanism. The optical zoom is performed at the lens and, although effective, is expensive to realize because of the cost of the lens. Instead, digital upscaling techniques may be used to increase the resolution of an area of interest, therefore creating the same effect as the optical zoom. The digital zoom block **151** may emulate the analog lens operation of zoom-in and zoom-out at a rate of change determined by the user. In a digital zoom operation the size of the picture remains the same.

The picture scaling block **152** may be implemented similarly to the digital zoom block **151**, except that the picture scaling block **152** generally operates on the entire picture, while the digital zoom generally operates on an area of interest. Scaling may be used in cases where the size of the picture needs to be modified in the frame of the original picture (e.g., to compose several pictures in one frame). Furthermore, picture scaling may be used to both increase and decrease the resolution on the input images in scaling ratios determined by the application and not necessarily directly by the user.

A resizing operation may be implemented that is a special case of picture scaling. The total number of samples used to represent the image are changed using digital resampling techniques (as opposed to the block **151** and the block **152** which generally preserve the original frame resolution). The block **153** may implement the resizing operation in a faster more efficient manner than the blocks **151** and **152**.

Referring to FIG. **4**, a more detailed block diagram is shown illustrating an example of a video/still picture Codec **132** implemented in accordance with one embodiment of the present invention. In one example, the block **132** may comprise a block (or circuit) **160**, a block (or circuit) **162**, a block (or circuit) **164**, a block (or circuit) **166** and a block (or circuit)

168. The block 160 may be implemented in one example, as a pre-processing module. The block 162 may be implemented, in one example, as an image analysis block. The block 164 may be implemented, in one example, as a still picture encoder/decoder (Codec). The block 166 may be implemented in one example, as a video encoder/decoder (Codec). In one example, the block 166 may be implemented as an H.264 compliant Codec. However, the block 166 may be compliant with other standards and/or proprietary compression schemes. The block 168 may be implemented, in one example, as a control block.

In one example, the signal PICTURE_DATA_IN may be presented to an input of the block 160. The block 160 may have a first output that may present a signal to an input of the block 162 and a second output that may present a signal (e.g., PICTURE_DATA) to an input of the block 164 and an input of the block 166. The information sent to the block 162 may include, but is not limited to, luminance picture data, chrominance picture data, low frequency and high frequency information on luminance and chrominance, object textures, object colors, object edges, temporal differences between pictures and motion estimated block vectors between pictures.

The block 162 may be configured to generate statistics and region data. The statistics and region data may be presented to a second input of the block 164, a second input of the block 166 and a first input of the block 168. The block 168 is implicitly connected to all of the blocks 160-166 in the block 132. The block 168 may be implemented as a top-level controller that manages operating modes (dictated by a user), quality requirements, and performance requirements. For example, the block 168 may be configured to manage rate control processes that may change the behavior of the block 166. The block 168 may be further configured to manage motion estimation processes of the block 166. The block 168 sets parameters and issues instructions for coordination of operations of the other blocks.

The block 166 may have a first output that may present the signal PICTURE_DATA_OUT, a second output that may present the signal BITSTREAM and a third output that may present information via the signal CODEC_DATA. In general, data that is not picture, image, bitstream, or any other image related data is processed by the system control block 168. The block 168 may have a second input that may receive the signal ISP_DATA, a third input that may receive the signal CTRL_INFO_IN, a first output that may present the signal CTRL_INFO_OUT and a second output that may present information via the signal CODEC_DATA. The block 168 may be configured to make decisions regarding system behavior. The block 168 generally executes the processes described below in connection with FIGS. 5-8.

Image data proceeding from the block 130 is generally preprocessed by the block 160 to improve compression quality (e.g., cropping, spatial scaling, color conversion and noise reduction). The image analysis module 162 is generally configured to use original and preprocessed data to produce a series of global (picture-based) and regional (specific area-based) statistics that represent image sequence characteristics.

The still picture encoder-decoder block 164 is generally configured to code still pictures. The still picture coding process is always part of a DSC since picture data compression is used to achieve efficient storage and transmission. In general, the still picture coding process 164 is included in the block 132 to illustrate that there is a unified coding approach. The image signal processing analysis data received from the block 130 (e.g., via the signal ISP_DATA) may be used to improve still picture quality. The process of using ISP analysis data to improve still picture quality is included in the coding parameter optimization phase illustrated in the flowcharts of FIGS. 5B, 6B and 7B.

In a conventional DSC, still picture coding is an external process to the ISP and greatly influenced by the capture efficiency/processing of the ISP. Typically, still picture compression is achieved by JPEG coding. By allowing both Codec analysis data and ISP analysis data to be fed into the still picture coding process, the present invention provides the ability to perceptually change the quantization parameters of selected areas of the picture.

The block 166 generally comprises several modules configured to process picture sequences in space and time to produce high quality compressed video bit streams. In one example, the block 166 may comprise a block (or circuit) 180, a block (or circuit) 182, a block (or circuit) 184, a block (or circuit) 186, a block (or circuit) 188, a block (or circuit) 190, a block (or circuit) 192 and a block (or circuit) 194. The block 180 may be implemented, in one example, as a motion estimation block. The block 182 may be implemented, in one example, as a transform and quantization block. The block 184 may be implemented, in one example, as a rate control block. The block 186 may be implemented, in one example, as a decoding block. The block 188 may be implemented, in one example, as a coding mode decision block. The block 190 may be implemented, in one example as a coding analysis block. The block 192 may be implemented, in one example, as a deblocking filter block. The block 194 may be implemented, in one example, as an entropy coder block. The blocks 180-194 may be implemented using conventional circuits and/or techniques modified in accordance with the teachings of the present disclosure. In one example, the block 186 may be omitted.

The blocks 180-194 are generally illustrative of processing modules in a hybrid entropy coding encoder. In general, the block 168 may be implemented to coordinate the functioning of the individual modules 180-194. Functioning of the encoder generally comprises control and data path connections configured to achieve compliance with a particular coding standard (or proprietary scheme).

In general, all of the modules 180-194 are influenced by the rate control block 184 since the block 184 keeps track of bit rate and quality requirements for particular operating modes set by higher level processes (e.g., user specifications and capture modes) controlled by the block 168. The motion estimation block 180 and the coding mode decision block 188 generally determine the most efficient encoding for the predicted data. The data may be predicted in the spatial direction (e.g., intra prediction) or the temporal direction (e.g., inter prediction). The transform and quantization block 182 is generally configured to convert the residual data (e.g., difference between actual and predicted data) into a domain most suitable for compression (e.g., the process of discarding perceptually irrelevant data). The quantization step is important for compression efficiency and one of the major controls for bit rate and quality constraints managed by the rate control block 184.

The decoding block 186 may be implemented as a standalone feature. The decoding block 186 generally duplicates the behavior of a decoder at the end of the transmission channel in order to avoid drift. Display functions are generally associated with the block 186 since in many cases it is desirable to monitor the encoded bit stream.

The deblocking block 192 is generically placed in the diagram to indicate the inclusion of some filtering during the compression process, whether inside the coding loop or out-

side of the coding loop. Compression quality is generally improved by the use of deblocking circuits.

The coding analysis block **190** may be configured to gather information about the coding process. The information gathered by the block **190** may be sent to the block **130** via the signal CODEC_DATA to improve processing. The type of data transmitted to the block **130** is outlined below.

In general, the block **132** may be configured to use data received from the block **130** via the signal ISP_DATA to change the behavior of the motion estimation (e.g., search window management, prediction modes, weighted prediction, etc.), quantization, rate control (e.g., to meet specific quality vs. bit rate targets), and noise reduction modules. In the system depicted in FIG. **2**, each of the blocks **130** and **132** is capable of generating analysis data relating to respective internal functioning and to the nature of the images processed. The image analysis module **156** generates the data and sends it forward to the block **132** via the signal ISP_DATA. The block **130** may be configured to generate data for use by the block **132** as described below. For example, information on lens distortion may be used to adjust coding parameters for optimal compression in areas that have less optical fidelity. Quantization may be coarser at the edges of the pictures where the distortion is more pronounced.

Hybrid entropy coding systems are partly based on detecting changes of motion in sequences. Therefore, there is always a need to estimate motion between pictures. Typical motion estimation (ME) engines detect changes based on some measure of correlation such as mean-squared error (MSE) or mean absolute error (MAE). These techniques tend to fail in situations where there is no real motion, but just changes in picture characteristics. Such cases may include sudden flashes of light (e.g., when a camera attempts to illuminate a subject for proper exposure); fade in and fade out (e.g., when exposure changes in order to produce the optimum exposure). In such cases, block **130** provides information about the current state of exposure and focus settings to block **132** so that motion estimation engines are set to operate under optimal conditions. For example, proper use of skip, direct and weighted prediction modes.

In state-of-the-art Codecs like H.264, there are inter-frame prediction modes and intra-frame prediction modes. Selecting the appropriate mode for efficient coding is a process that involves multiple tries of intra and inter coding modes to find the best performance; or a-priori knowledge of the material to select only the appropriate coding modes.

Quantization is the process of converting the transform coefficients to fixed point representation. As with any quantization process, small coefficients will be quantized to zero value if the quantization parameter is large (coarse quantization parameter) and therefore, those coefficients can be regarded as non-critical for picture coding fidelity. However, there are many cases where particular picture features are important for achieving good fidelity and therefore should not be quantized coarsely, but rather with a finer (smaller) quantization step. Knowledge of image and sequence characteristics help in determining those areas that should be quantized more carefully than other to achieve good picture fidelity.

Achieving acceptable fidelity and high compression is a common optimization problem since fidelity and high compression constitute conflicting goals. In particular, excellent image quality may lead to low (inefficient) compression, while high (efficient) compression tends to produce poor image quality. The rate control block **184** is generally configured to provide a good compromise between picture quality and compression efficiency. A-priori knowledge of picture and sequence characteristics can help manage the optimization task described above.

Pictures devoid of random noise are easier to encode because there is less irrelevant high frequency information. However, the unabridged application of noise reduction techniques to images in both the spatial and temporal domain can lead to soft and flat pictures that do not represent real scenery. Data regarding the capture process may be obtained in order to make better decisions as to the amount of noise reduction needed and the coding modes needed for faithful reproduction at all target bit rates.

The block **132** may be configured to send adjustments to exposure parameters of the block **130** via the signal CODEC_DATA based on compression efficiency and quality performance. The exposure settings may be adjusted so that certain areas of the picture are optimally encoded. The block **132** may identify overly dark and overly bright areas of the picture for better exposure and adjust the quantization approach so that those areas are assigned enough bits for proper coding. The picture gain of the block **130** may be adjusted via the signal CODEC_DATA to improve motion estimation performance and better rate control when the block **130** is in a dynamic configuration phase (e.g., auto exposure, auto focus, etc.). Fast changes in the picture gain due to auto exposure hunting may be minimized under very low bit rate conditions (e.g., so as not to overwhelm the encoding process).

The block **132** may be configured to send adjustments to the focus parameters of the block **130** via the signal CODEC_DATA based on compression efficiency and quality performance. For example, focus may be slightly de-tuned under extremely low bit rate requirements in order to improve compression performance. Focus dynamics of the block **130** may be adjusted during the auto-focus phase in order to improve motion estimation performance and rate control (e.g., sudden focus changes may be minimized in order to reduce sudden bit production during encoding). Adjustments to the noise reduction process of the block **130** may be sent via the signal CODEC_DATA based on compression efficiency and quality performance. For example, under low lighting conditions and low motion, the block **132** may send commands for increased noise reduction since performance will not be affected by high motion.

Global statistics data may be obtained by the preprocessing module **160** after analyzing each picture as a whole (e.g., with no particular area of each picture having more importance than another). Region-based statistics data may be obtained by classifying areas of the picture with certain feature characteristics such as amount of low/high frequency texture, amount of edges, amount of motion detected, amount of moving edges, type and amount of a specific color, etc.

In contrast with control data sent via the signal CODEC_DATA to the block **130**, the raw global and region-based statistics data may be sent as information to be processed and acted upon at the discretion of the respective image signal processing procedures of the block **130** (e.g., no declarative command actions are associated with the data). For example, size and strengths of regions in the picture classified according to a set of features (e.g., texture, edges, motion, color, etc.) may be provided to the block **130** to evaluate the amount and relevance of featured regions in the picture in both the spatial and temporal domains. Information regarding areas of low/high texture in the picture may be sent via the signal CODEC_DATA to the block **130**. The block **130** may use the texture information for auto-focus processing by

evaluating the amount of low/high frequency in the picture as a result of changes of focal length by the focusing circuits.

Information regarding areas of the picture with strong edge components may be sent from the block **132** to the block **130** via the signal CODEC_DATA. The block **130** uses the edge information for auto-focusing whereby the amount of edges in the picture increase as the picture comes into focus. This mechanism may be used along with motion information to determine the quality of auto-focus data. Furthermore, edge information can be used in demosaicing procedures that rely on edge information to adapt the Bayer conversion process.

The block **132** may send information regarding areas of low/high motion in the picture, including moving edges, to the block **130** via the signal CODEC_DATA. The block **130** may use the motion information for auto-focusing (e.g., to determine quiescent areas of the picture upon which to reliably focus), auto-exposure (e.g., by increasing exposure time when pictures are non-moving) and noise reduction (e.g., by applying stronger noise reduction to non-moving areas).

The block **132** may send further information concerning areas of the picture featuring a specific color to the block **130** via the signal CODEC_DATA. In one example, the block **132** may be configured to operate in the YUV domain to facilitate detection of areas of specific colors. The color information may be used to identify human faces to help reduce red-eye effects (which also depend upon identification of red areas of the picture). The block **130** may use the color information to reduce the red-eye effects when using a flash.

In addition to the control data and global picture and region-based statistics, the block **132** may be configured to send statistics derived from analysis of the encoding process via the signal CODEC_DATA using a mutually agreed protocol for data interchange. For example, motion vector information may be transmitted to the block **130** to implement/improve image stabilization processes. The block **130** may use the motion vector information to obtain detailed motion in specific areas of the picture for improved motion-adaptive procedures under low lighting conditions (e.g., auto exposure, auto-focus, noise reduction, image stabilization). In addition to the motion vector information, block distortion information may be sent to the block **130**. For example, the block distortion information may include an indication of the coding fidelity at a given bit rate. Distortion per block at the edges of the picture analyzed in the temporal domain may be an indication of fixed lens aberrations that may be corrected by the block **130**.

In addition to sending the above information to the block **130** via the signal CODEC_DATA, information generated in the block **132** may also be sent to the block **164**. The signal CODEC_DATA generally represents all information gathered by block **132** in the process of preprocessing and encoding the moving picture data. In one example, the block **164** may be implemented as a JPEG or JPEG2000 still picture codec. The signal CODEC_DATA may contain information to improve performance of block **130** as indicated on the left side of the flow charts in FIGS. **5-8**. For example, in FIG. **8**, when the flash is active, the block **130** may use information in the signal CODEC_DATA to change the flash intensity and duration.

In one example, the block **164** may be implemented as part of the block **132**. However, in other implementations the block **164** may be implemented as part of the block **130**, or as a separate system altogether. In one example, the block **164** may be a separate system implemented with a general DSP Processor.

Motion information concerning areas of low/high motion in the picture, including moving edges may be used to improve perceptual coding. For example, areas of low motion may be coded with a fine quantization step in order to reproduce them faithfully. On the other hand, areas of high motion are presumably blurred and, depending on the ISP parameters received, may be coded with very few bits if over/under exposed or adversely affected by low shutter speeds. Color information about areas of the picture featuring a specific color may be used to identify human faces for perceptual coding. The quantization step may be much lower in these areas in order to reproduce faces more faithfully.

The block **166** may be configured to send statistics derived from analysis of the encoding process to the block **164** to steer perceptual processing performed by the block **164**. In one example, the statistics sent to the block **164** may include a number of bits coded per block, block distortion, block coded residual variance and number of quantized transform coefficients with zero value per picture. The number of bits coded per block may be sent to the block **164** to give an indication of coding efficiency per picture. The block **164** may collect the number of bits coded per block information for every picture in order to improve internal processing procedures and also to make decisions regarding the effect of the internal processing procedures on compression efficiency.

The block distortion information may be sent to the block **164** as an indication of the coding fidelity at a given bit rate; together with the number of bits per block, the block distortion information may be used to evaluate the total efficiency of the encoding process in terms of the compromises between distortion and rate. For example, the block **166** generally uses the number of bits per block and block distortion values to tune rate-control processes. The block coded residual variance data are an indication of coding efficiency in both the spatial and temporal domains as a result of coding decisions in the block **166** due to motion estimation and quantization. The block **164** may use the block coded residual variance data to adapt internal processing methods by allocating less bits to areas where the residual is low (e.g., assuming that the areas where the residual is low can be coded efficiently with a higher quantization step). The data regarding the number of quantized transform coefficients with zero value per picture are an indication of compression efficiency with respect to picture/sequence complexity. The information may be used by the block **164** in the same manner as the number of bits per block.

Referring to FIGS. **5**(A-B), flow diagrams are shown illustrating an example of a zoom operation of the block **130** (FIG. **5**A) and the block **132** (FIG. **5**B) of FIG. **2**. In one example, a zoom procedure **200** for the block **130** may comprise a state **201**, a state **202**, a state **204**, a state **206** and a state **208**. The zoom procedure **200** for the block **130** may start in the state **201** and move to the state **202**. The state **202** may comprise a decision state. In the state **202**, the process **200** may determine whether a zoom function is active. When the zoom function is active, the process **200** may move to the state **204**. If the zoom function is not active, the process **200** may move to the state **208**.

In the state **204**, the process **200** may determine whether the zoom operation is zooming in (e.g., enlarging or magnifying an object of interest). The operation of zooming in may also be referred to as the zoom moving in. When the zoom operation is not zooming in, the process **200** may transition to the state **206**. When the zoom operation is zooming in, the process **200** may determine the speed (e.g., rate of change of magnification) at which the zoom is moving in and transition to the state **208**. In the state **206**, the process **200** may determine whether the zoom operation is zooming out (e.g., reducing an apparent size or de-magnifying an object of interest).

The operation of zooming out may also be referred to as the zoom moving out. When the zoom operation is not zooming out, the process **200** may collect information regarding the current magnification and transition to the state **208**. When the zoom operation is zooming out, the process **200** may determine a speed (e.g., rate of change of the magnification) at which the zoom is moving out and transition to the state **208**. In the state **208**, the information (e.g., magnification and/or rate of change of the magnification) may be collected and sent to the block **132**.

A zoom procedure **210** for the block **132** may comprise a state **212**, a state **214**, a state **216**, a state **218**, a state **220**, a state **222**, a state **224** and a state **226**. The state **212** may be implemented as a start state. The state **214** may be implemented as a zoom setting collection state. The state **216** may be implemented as a decision state. The state **218** may be implemented as a decision state. The state **220** may be implemented as a decision state. The state **222** may be implemented as an adjustment state. The state **224** may be implemented as an optimization state. The state **226** may be implemented as a completion state.

The process **210** may begin in the state **212** and transition to the state **214**. In the state **214**, the process **210** generally obtains zoom settings from the process **200**. When the zoom settings have been obtained, the process **210** generally transitions to the state **216**. In the state **216**, the process **210** determines whether the zoom is active. When the zoom is not active, the process **210** generally transitions to the state **226** and terminates. When the zoom is active, the process **210** generally transitions to the state **218**.

In the state **218**, the process **210** determines whether the zoom operation is zooming in. When the zoom operation is not zooming in, the process **210** generally transitions to the state **220**. When the zoom operation is zooming in, the process **210** generally obtains the speed (e.g., rate of change of magnification) at which the zoom is moving in from the process **200** and transitions to the state **222**. In the state **222**, the process **210** generally adjusts motion estimation and/or rate control parameters and transitions to the state **226** for completion.

In the state **220**, the process **210** generally determines whether the zoom operation is zooming out. When the zoom operation is not zooming out, the process **210** generally obtains the magnification determined in the process **200** and transitions to the state **224**. When the zoom operation is zooming out, the process **210** generally obtains the speed (e.g., rate of change of magnification) at which the zoom is moving out determined in the process **200** and transitions to the state **222**. In the state **224**, the process **210** generally optimizes coding parameters based on the magnification determined and transitions to the state **226**.

The block **132** may be configured to modify motion estimation (ME) and mode decision processes specifically for zooming-in or zooming-out portions of the input sequence. Under fast zoom situations, the block **132** may be configured to adjust rate control processes for efficient coding. When zooming is in a quiescent state, picture parameters may be optimized for a current magnification factor. For example, under high magnification unstable camera hold may be a problem and therefore rate control should be set to ignore quick changes. The process represented by the optimize coding parameters block **224** generally involves non-dynamic changes to the processes indicated in adjacent adjust motion estimation and adjust rate control blocks **222**.

During zoom, the motion estimation (ME) process (e.g., represented by the block **180** in FIG. **4**) may produce unreliable motion vectors because practical ME implementations are only valid for translational motion. A low zoom speed may be characterized, for example, by a global picture change of less than 5% per picture. A high zoom speed may be defined, for example, as a global picture change of more than 10% per picture. Changes in between these values may be considered to be medium zoom speeds. During a zoom process, ME may be set to favor smaller vectors as long as the distortion does not exceed a certain threshold. The threshold may be used to guarantee that when a large vector is found with very small distortion, the large vector will still be chosen over a smaller vector. However, for the great majority of cases, the smaller vector may be chosen. The threshold may be set, in one example, as a mean squared error (MSE) of 10% for the current block being processed.

During the zoom process, ME may favor co-located motion prediction. For example, the starting search location may be the same position of the current block in the reference picture and not the predicted location. The co-located ME mode of operation may be used generally in the special cases, however, the Codec may also try other starting search locations. For example, the Codec may try one from a PMV (predicted motion vector). The PMV is a standard feature of H.264. The PMV predicts the nature of the motion vector for the block being processed using a normative algorithm in order to improve efficiency of the ME process. However, sometimes under special circumstances, starting the search at location zero (the co-located location of the block under processing) may make more sense.

For low to medium zoom speeds, ME may be set to favor small blocks in areas that contain edges. For medium to high zoom speeds, ME may favor intra coding modes when the distortion is larger than the chosen inter coding mode by a small percentage. In one example, the percentage value may be set to 5%. Setting the percentage value at 5% may ensure that the coded information is not biased by incorrect, but efficient, motion vectors. For high zoom speeds small vectors may be favored. When the frequency content of the block is low (e.g., less than 3%), intra coding modes may be chosen. For all zoom cases, the ME capability for reference picture resampling is generally enabled. For example, the reference data may actually be a resampled (scaled) version of the original reference frames. The ME control process determines the direction to resample (e.g., up or down) depending on the direction of zoom (e.g., in or out, respectively).

When the block **130** controls the camera to zoom out (e.g., objects in the picture are getting smaller because of increased field of view), the ME block may be biased to produce small vectors and intra blodks at the edges of the picture. For example, new details generally enter the picture at the edges, so there is no point in trying to perform a motion estimation match for large search areas. In cases where the zoom varies between zooming-in and zooming-out (e.g., when a user is trying to achieve optimal picture cropping), ME may be set to favor multiple reference pictures. The use of multiple reference pictures may be justified because in the process of zooming in and out the pictures will generally be very similar and the likelihood of the motion estimation encountering a good match may be very high.

The modification of motion estimation based on the zoom function described above may be summarized with the following example pseudo code:

```
If (zoom active) {
  RefPictResampling = TRUE
  If (MBFreq == LOW)
    IntraCodingBias = TRUE
```

-continued

```
If (ZoomOut) {
   MEmode = small motion vectors
   IntraCoding @ Edge of Picture
}
If (ZoomOut AND ZoomIn)
   MultipleRefPictures = TRUE
If (ZoomSpeed == HIGH)
   MEmode = small motion vectors
If (ZoomSpeed == HIGH OR ZoomSpeed == MED)
   If (Distortion > Threshold)
      IntraCodingBias = TRUE
If (ZoomSpeed == MED OR ZoomSpeed == LOW)
   MEmode = small blocks
}
```

Referring to FIGS. **6**(A-B), flow diagrams are shown illustrating an example of an auto exposure process for the block **130** (FIG. **6**A) and the block **132** (FIG. **6**B). The process **230** may comprise state **232**, a decision state **234** and a state **236**. In the state **232**, the block **130** generally begins the auto exposure procedure and transitions to the state **234**. In the state **234**, the process **230** determines whether or not the auto exposure mode is active. If the auto exposure mode is active, the process **230** transitions to the state **236** and collects information regarding F-stop and shutter speed. When the auto exposure mode is not active, the process **230** generally transitions to the state **236** and no data is collected or sent. During the auto exposure procedure **240**, the block **132** generally enters the state **242** to begin the procedure for auto exposure. The process **240** generally transitions to the state **244**. In the state **244**, the process **240** determines whether the automatic exposure mode is active. If the automatic exposure mode is not active, the process **240** generally transitions to the state **246**. If the auto exposure mode is active, the process **244** generally transitions to the state **248**. In the state **246**, the process **240** generally optimizes coding parameters and transitions to the state **252**. In the state **248**, the process **240** generally adjusts rate control parameters, adjusts quantization parameters and adjusts motion estimation parameters. The adjustments to the rate control, quantization and motion estimation may be performed taking into account statistics received from the Codec (e.g., the block **252**). Once the adjustments of rate control, quantization and motion estimation have been completed, the process **240** may transition to the state **250** for completion.

When the block **130** adjusts exposure settings dynamically, the block **132** may be configured to adjust quantization and rate control for efficient coding. Change in global exposure settings may be handled with weighted prediction modes by the block **132** therefore improving coding rates. Dynamic range changes to selected areas of the picture may be coded with more/less bits depending on the relevance of the regions in terms of psycho-visual measures (e.g., more important areas may be allocated more bits). Both data from the block **130** and statistics from the block **132** may be used to improve the final coding efficiency of the output bit streams. In general, the optimized coding parameters block **246** generally represents non-dynamic changes to the processes indicated in the adjacent rate control, quantization and ME blocks **248**.

The automatic exposure (AE) process generally changes the luminance of the pictures with the aim of reaching an optimum exposure setting for best picture quality. In the context of using an ISP with a Codec, achieving the best picture quality that yields the best compression quality is desirable. For example, adjustments may be made such that the range of luminance for the pictures to be encoded is linear in the midtone range, while de-emphasizing the dark and light areas of the picture. Such a transfer function is sometimes referred to as an "S-shaped" curve because the graphical representation resembles a letter S when the brightness input is represented on the abscissa axis and the bright output is represented on the ordinate axis.

When AE is active, ME may favor prediction from co-located block locations. The ME may also favor smaller motion vectors. Rate of change and amount of change in pictures is generally provided by the Codec Statistics (e.g., the block **252**). ME may favor Weighted Prediction modes when a substantially linear luminance change is detected. For example, a linear change may be described by a generic linear equation $Y=mx+b$, where m is referred to as the slope (or rate of change) and b represents a generic offset. Given a set of values $\{Y0, Y1, Y2, Yn-1, Yn\}$, if the ratio of two consecutive values (e.g., $Yi/Y(i-1)$) is constant, then the change may be said to be linear. In one example, several luma(Y) values may be sampled and the ratio between the values determined. If the variation in the ratio for a number of successive samples (e.g., 15 samples) is within a predetermined range (e.g., 1% of each other), the change may be said to be substantially linear. Since instantaneous exposure changes translate into global luminance changes in each picture, a weighted factor may be computed from changes made by the automatic exposure process in subsequent pictures. Weighted prediction factors (e.g., offset and multiplicative factors) maybe controlled based on exposure times when auto-exposure is active.

When small global motion changes are reported in the Codec Statistics, ME control block **180** may evaluate the efficiency of Long Term Picture coding, and when global changes between pictures amount to, for example, less than 5% of total brightness, pictures designated as Long Term may be used for prediction instead of the current picture. When a Codec does not have the Long Term Picture or Weighted Prediction facility, the current picture may be Skipped in its entirety while the AE mechanism is active. Using Skip ME modes when AE is active and there is low global motion in the pictures may save bits by repeating pictures at a lower rate since faithful reproduction of the AE process is not critical for video content.

The modification of motion estimation based on automatic exposure considerations described above may be summarized with the following example pseudo code:

```
If (AE active) {
   MEmode = collocated Macroblock
   MEmode = small motion vectors
   If (GlobalMotion == LOW)
      LongTermPicture = ENABLE, or SkipCurrentPict = ENABLE
   If (LinearLumaChange)
      MEmode = weighted prediction
}
```

Referring to FIGS. **7**(A-B), flow diagrams are shown illustrating an example of an auto-focus process for the block **130** (FIG. **7**A) and the block **132** (FIG. **7**B). The process **260** may comprise state **262**, a decision state **264** and a state **266**. In the state **262**, the block **130** generally begins the auto-focus procedure and transitions to the state **264**. In the state **264**, the process **260** determines whether or not the auto-focus mode is active. When the auto-focus mode is active, the process **260** transitions to the state **266** and collects information regarding focus speed. When the auto-focus mode is not active, the process **260** generally transitions to the state **266** and collects information regarding focal length.

During the auto-focus procedure, the block **132** generally performs a process **270**. The process **270** generally comprises a state **272**, a state **274**, a state **276**, a state **278** and a state **280**. The process **270** generally entered the state **272** to begin the procedure for auto-focus. The process **270** generally transitions to the state **274**. In the state **274**, the process **270** determines whether the automatic focus mode is active. When the automatic focus mode is not active, the process **270** generally transitions to the state **276**. When the auto-focus mode is active, the process **270** generally transitions to the state **278**.

In the state **276**, the process **270** generally optimizes coding parameters and transitions to the state **280**. In one example, focal length information may be used to optimize the coding parameters. In the state **278**, the process **270** generally adjusts rate control parameters, adjusts quantization parameters and adjusts motion estimation parameters. The adjustments to the rate control, quantization and motion estimation may be performed taking into account speed information (e.g., rate of change of focus). Once the adjustments of rate control, quantization and motion estimation have been completed, the process **270** may transition to the state **280** for completion.

During focusing adjustments, the block **130** may be configured to change focal length dynamically while the block **132** may be configured to adjust rate control for optimal quality (e.g., using lower bit rates to save bits for later more complex scenes). During the focusing operation the block **132** generally uses low bit rates since the video is not yet optimized for recording (blurry). The block **132** may be configured to modify motion estimation processes to work more efficiently during focal changes. The optimize coding parameters block **276** generally represents non-dynamic changes to the processes indicated in the adjacent ME, Quantization and Rate Control blocks **278**.

When the automatic focus operation is being performed, the sharpness in the pictures may fluctuate. Information passed from the process **260** (e.g., from the ISP **130**) to the process **270** (e.g., to the Codec **132**) may indicate the fluctuation in picture sharpness. The rate of change in focus is directly related to the rate of change of global high frequency in the pictures in a video sequence to be coded. ME may emphasize small motion co-located motion vectors when AF is active. As in the AE case described above, when the global motion is low and the AF is in hunting mode (e.g., going back and forth between a specific focus point), ME may favor Skip and Long Term Picture (LTP) prediction modes. When motion between pictures is very low, using Skip mode to code may be more efficient. For example, for every ten pictures with low motion one picture may be selected as a Long Term Picture to be used as a reference if a specific focus point is revisited during the hunting operation. Both Skip and LTP modes may be used during prolonged hunting auto-focus operations. When the ISP is hunting for optimal focus, the ME block may be set to favor multiple reference pictures using small search ranges.

In general, the present invention uses the small co-located motion search ranges very consistently depending on the type of operation being performed by the ISP. In one example, small motion search range may be defined as 3% of the picture size (e.g., for a 720×480 picture the search range may be around 16×16 in horizontal x vertical directions).

The modification of motion estimation based on automatic focus consideration described above may be summarized in the following TABLE 1:

TABLE 1

	Small MV	Co-located MB	Skip Mode	Long Term Picture	Multiple Reference
AF Active	X	X			
AF Hunting	X				X
AF Hunting and LowGlobalMotion	X		X	X	X

Referring to FIGS. **8**(A-B), flow diagrams are shown illustrating an example of flash process **300** for the block **130** (FIG. **8**A) and a flash process **310** for the block **132** (FIG. **8**B). The process **300** may comprise state **302**, a decision state **304** and a state **306**. In the state **302**, the block **130** generally begins the flash procedure and transitions to the state **304**. In the state **304**, the process **300** determines whether or not the flash is active. When the flash is active, the process **300** transitions to the state **306** and collects information regarding flash intensity and duration. When the flash is not active, the process **300** generally transitions to the state **306** and no data is collected or sent.

The process **310** generally comprises a state **312**, a state **314**, a state **316**, and a state **318**. During the flash procedure **310**, the block **132** generally enters the state **312** to begin the procedure for flash compensation. The process **310** generally transitions to the state **314**. In the state **314**, the process **310** determines whether the flash is active. When the flash is active, the process **310** generally transitions to the state **316**. When the flash is not active, the process **310** generally transitions to the state **318**.

In the state **316**, the process **310** generally adjusts rate control parameters, adjusts quantization parameters and adjusts motion estimation parameters. The adjustments to the rate control, quantization and motion estimation parameters may be performed taking into account information about flash intensity an/or duration. Once the adjustments of the rate control, quantization and motion estimation parameters have been completed, the process **310** may transition to the state **318** for completion.

Sudden flash pictures can overwhelm a common video encoder. By transmitting flash intensity and duration information from the block **130** to the block **132**, the block **132** may adjust motion estimation and rate control to deal with the sudden change in pictures. For example, at the flash transitions, the block **132** may insert new Intra prediction frames. In H.264 coding, a Long Term Picture may also be generated at the flash occurrence.

The ME is generally in the same mode of operation as in the AE case above, except as follows. When the ISP computes the intensity of the flash discharge for optimal image quality, it is possible that external events may cause an over-exposed or under-exposed picture. The image analysis process of the Codec **132** (e.g., the block **162** in FIG. **4**) may be used to measure the total brightness of each picture. When a sudden trend is detected (e.g., when the average picture brightness from one picture to another picture increases by more than 50%), the ME may be operated in the same mode as when AE is active. However, when there is a sudden increase in brightness (e.g., over-exposure), the ME may use Long Term Prediction (LTP). For example, the ME may make use of a previously generated picture to be used as a long term reference. In one example, the ME may use ZERO motion vectors (since when over-exposure occurs all vectors will be invalid). When LTP is not available, an Intra frame may be inserted and ME may be essentially disabled.

The terms "favor", "emphasize" or "bias" in some of the modes mentioned in this disclosure generally refer to a prob-

ability of the mode being chosen being increased. As would be apparent to those skilled in the relevant art(s), increasing the probability of the mode being chosen may be accomplished in a variety of ways, including, but not limited to, the assignment of likelihood parameters using a variety of implementation choices (e.g., look-up tables, case statements in declarative or imperative computer languages, multivaried logic, etc.). In a rate-distortion optimized encoder, the encoder may favor small vectors by using a large langrangian parameter during motion vector cost calculation. For example, a motion vector cost (e.g., MVCOST) may be determined by a equation such as MVCOST=lambda*rate+distortion, where rate represents the rate or size of the motion vectors, and distortion comprises SAD, SSE, MAD or MAE. The encoder may favor skip and/or co-located prediction modes by using a larger lambda for those modes during motion estimation and mode decision. The term camera, as used herein, generally refers to any digital image acquisition and processing device. In one example, the digital image acquisition and processing device may be configured to process a sequence of pictures constituting a video stream. In general, the present invention may be applied to devices using a variety of sensors including, but not limited to, visible light, infrared (IR), ultraviolet (UV), x-ray, ultrasound, etc.

As would be apparent to those skilled in the relevant art(s), the signals illustrated in FIGS. **1-4** represent logical data flows. The logical data flows are generally representative of physical data transferred between the respective blocks by, for example, address, data, and control signals and/or busses. The system **100** may be implemented in hardware, software or a combination of hardware and software according to the teachings of the present disclosure, as would be apparent to those skilled in the relevant art(s).

The functions performed by the flow diagrams of FIGS. **5-8** may be implemented using a conventional general purpose digital computer programmed according to the teachings of the present specification, as will be apparent to those skilled in the relevant art(s). The pseudo code presented above is provided as an exemplary description and should be qualified with the respective textual description for complete understanding. Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will also be apparent to those skilled in the relevant art(s).

The present invention may also be implemented by the preparation of ASICs, FPGAs, or by interconnecting an appropriate network of conventional component circuits, as is described herein, modifications of which will be readily apparent to those skilled in the art(s).

The present invention thus may also include a computer product which may be a storage medium including instructions which can be used to program a computer to perform a process in accordance with the present invention. The storage medium can include, but is not limited to, any type of disk including floppy disk, optical disk, CD-ROM, magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, Flash memory, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

While the invention has been particularly shown and described with reference to the preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made without departing from the spirit and scope of the invention.

The invention claimed is:

1. A camera comprising:

a first circuit configured to receive an input signal and perform image signal processing on said input signal using camera setting information and encoding related information;

a second circuit configured to encode image data using image signal processing related information and said camera setting information, the second circuit comprising a still picture encoder-decoder block and a video encoder-decoder block for processing still pictures and picture sequences, respectively; and

wherein (i) said first circuit is further configured to pass said image data and said image signal processing related information to said second circuit, and (ii) said second circuit is further configured to pass said encoding related information to said first circuit and modify one or more motion estimation processes based upon said image signal processing related information from said first circuit.

2. The camera according to claim **1**, wherein said first and said second circuits are integrated in a single integrated circuit.

3. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor small motion vectors as long as distortion is below a predetermined threshold when a zoom process is active.

4. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor co-located motion prediction when a zoom process is active.

5. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor small blocks in areas containing edges for medium zoom speeds.

6. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor intra coding modes to reduce distortion for zoom speeds ranging from medium to high.

7. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to enable reference picture resampling with a direction of the resampling being determined by a direction of zoom.

8. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to enable reference picture resampling with a resampling ratio being determined by a zoom ratio.

9. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor small motion vectors and intra blocks at edges of a picture while zooming out.

10. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor multiple reference pictures when zoom is varied between zoom-in and zoom-out.

11. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor co-located motion prediction and small motion vectors when an automatic exposure process is active.

12. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor a weighted prediction mode when a substantially linear luminance change is detected.

13. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to control weighted prediction factors based on exposure times when an auto-exposure process is active.

14. The camera according to claim **13**, wherein weighted prediction factors comprise one or more factors selected from the group consisting of offset factors and multiplicative factors.

15. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes such that when small global motion changes are reported in coding statistics the efficiency of Long Term Picture coding is evaluated and pictures designated as Long Term are used for prediction instead of a current picture when global changes between pictures are less than a predetermined percentage of total brightness.

16. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes such that when small global motion changes are reported in coding statistics and automatic exposure is active either a current picture is skipped or skip mode is favored.

17. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes such that small motion co-located motion vectors are emphasized when an automatic focus process is active.

18. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor Skip or Long Term Picture prediction modes when global motion is low and an automatic focus process is in a hunting mode.

19. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes to favor multiple reference pictures using small search ranges when an automatic focus process of said first circuit is hunting for optimal focus.

20. The camera according to claim **1**, wherein said second circuit modifies said one or more motion estimation processes such that (i) when a trend in average picture brightness from one picture to another picture exceeds a predetermined threshold the motion estimation processes operate as though automatic exposure is active and (ii) when a sudden increase in brightness or over-exposure is detected the motion estimation processes use either a Long Term Picture prediction mode, zero motion vectors or insertion of an intra frame.

21. A camera comprising:

means for receiving an input signal and performing image signal processing on said input signal using encoding related information and camera settings information;

means for performing image encoding using image signal processing related information and said camera settings information, and said camera setting information, the second circuit comprising a still picture encoder-decoder block and a video encoder-decoder block for processing still pictures and picture sequences, respectively; and

wherein (i) the image signal processing means is further configured to pass the image signal processing related information to the image encoding means, and (ii) the image encoding means is further configured to pass the encoding related information to the image signal processing means and modify one or more motion estimation processes based upon the information from the image signal processing means.

22. A method for control of motion estimation comprising the steps of:

performing an image signal processing operation using camera settings information and encoding related information within a camera; and

performing an image encoding operation using image signal processing related information and said camera settings information within the camera, and said camera setting information, the second circuit comprising a still picture encoder-decoder block and a video encoder-decoder block for processing still pictures and picture sequences, respectively; and

wherein (i) the image signal processing operation is further configured to pass the image signal processing related information to the image encoding operation, and (ii) the image encoding operation is further configured to pass the encoding related information to the image signal processing operation and modify one or more motion estimation processes based upon the image signal processing related information.

23. The method according to claim **20**, wherein:

when zoom is active said one or more motion estimation processes are modified to favor one or more criteria selected from the group consisting of (i) small motion vectors with distortion below a predetermined threshold, (ii) co-located motion vectors, (iii) small blocks in areas containing edges for medium zoom speeds, (iv) intra coding modes to reduce distortion for zoom speeds ranging from medium to high, (v) reference picture resampling with a resampling direction determined by a direction of zoom, (vi) reference picture resampling with a resampling ratio determined by a ratio of zoom, (vii) small motion vectors and intra blocks at edges of a picture while zooming out and (viii) multiple reference pictures when zoom is varied between zooming-in and zooming-out;

when automatic exposure is active said one or more motion estimation processes are modified to favor one or more criteria selected from the group consisting of (i) co-located motion prediction and small motion vectors, (ii) weighted prediction in response to detection of a substantially linear luminance change, (iii) controlling weighted prediction factors based on exposure time, (iv) controlling offset and multiplicative factors based on exposure time, (v) skipping a current picture when small global motion changes are reported in coding statistics, (vi) favoring skip mode when small global motion changes are reported in coding statistics and (vii) using a Long Term Picture for prediction instead of a current picture when global changes between pictures are less than a predetermined percentage of total brightness; and

when automatic focus is active said one or more motion estimation processes are modified to favor one or more criteria selected from the group consisting of (i) small motion co-located motion vectors, (ii) Skip or Long Term Picture prediction when global motion is low and the automatic focus is in a hunting mode and (iii) multiple reference pictures using small search ranges when the automatic focus is hunting for optimal focus.

* * * * *